Title (en)

MULTIMODAL IMAGE CLASSIFIER USING TEXTUAL AND VISUAL EMBEDDINGS

Title (de)

MULTIMODALER BILDKLASSIFIZIERER MIT TEXTUELLEN UND VISUELLEN EINBETTUNGEN

Title (fr)

CLASSIFICATEUR D'IMAGES MULTIMODAL UTILISANT DES INTÉGRATIONS TEXTUELLES ET VISUELLES

Publication

EP 3791322 A1 20210317 (EN)

Application

EP 19818391 A 20191118

Priority

- US 201862768701 P 20181116
- US 2019061950 W 20191118

Abstract (en)

[origin: WO2020102797A1] Methods, systems, and apparatus, including computer programs encoded on a computer storage medium, for realizing a multimodal image classifier. In an aspect, a method includes, for each image of a plurality of images: processing the image by a textual generator model to obtain a set of phrases that are descriptive of the content of the image, wherein each phrase is one or more terms, processing the set of phrases by a textual embedding model to obtain an embedding of predicted text for the image, and processing the image using an image embedding model to obtain an embedding of image pixels of the image. Then a multimodal image classifier is trained on the embeddings of predicted text for the images and the embeddings of image pixels for the images to produce, as output, labels of an output taxonomy to classify an image based on the image as input.

IPC 8 full level

G06K 9/62 (2006.01)

CPC (source: EP US)

G06F 18/214 (2023.01 - US); **G06F 18/24** (2023.01 - EP US); **G06F 18/24147** (2023.01 - US); **G06V 10/774** (2022.01 - EP); **G06V 10/82** (2022.01 - EP)

Designated contracting state (EPC)

AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

Designated extension state (EPC)

BA ME

DOCDB simple family (publication)

WO 2020102797 A1 20200522; CN 112204575 A 20210108; EP 3791322 A1 20210317; US 11907337 B2 20240220; US 2021264203 A1 20210826; US 2024143700 A1 20240502

DOCDB simple family (application)

US 2019061950 W 20191118; CN 201980036217 A 20191118; EP 19818391 A 20191118; US 201917046313 A 20191118; US 202418409411 A 20240110